

## 비관측요인모형을 이용한 한국의 국내총생산 분석<sup>†</sup>

성병찬<sup>1</sup> · 이승경<sup>2</sup>

<sup>1</sup>중앙대학교 응용통계학과 · <sup>2</sup>중앙대학교 통계학과

접수 2011년 7월 13일, 수정 2011년 8월 7일, 게재확정 2011년 8월 15일

### 요약

본 논문에서는 비관측요인모형을 이용하여 한국의 국내총생산 시계열 자료를 분석한다. 이 모형이 확률적 및 결정적 요인들을 모두 포괄할 수 있다는 점을 이용하여, 보다 다양한 형태로 시계열 자료의 모형화를 시도하였으며, 지수평활법 및 박스-젠킨스의 ARIMA모형과 예측력을 비교하였다. 국내총생산 자료에 대한 2년간의 미래 예측에서 비관측요인모형이 보다 우수함을 보인다.

주요어: 구조적 시계열 모형, 상태공간모형, 확률적 추세.

### 1. 서론

시계열 자료 분석에 있어서 가장 어려운 문제 중 하나는, 결정적 추세 (deterministic trends) 또는 확률적 추세 (stochastic trends)를 선택하는 것이다 (Diebold와 Senhadji, 1996; Rao, 2010). 두 가지 추세를 선택하는 일반적인 기준은 없으며, 결정적 추세가 더 널리 사용되고 있고 확률적 추세는 비교적 덜 알려져 있다. 예를 들면, 단위근 검정 (unit root tests) 및 공적분 검정 (cointegration tests)에서 고정된 (fixed or constant) 수준 (level) 및 기울기 (slope)를 가지는 시계열의 추세가 모형에 주로 도입된다. 한편, Perron (1989)와 Perron과 Wada (2005)는 변화점 (change point)을 가지는 결정적 추세를 고려하였다. 이에 반하여, Harvey (1997)는 비교적 짧은 기간의 시계열에서는 고정된 직선의 결정적 추세가 시계열의 추세를 적절하게 반영할 수 없으며 너무 제약적임을 지적하였다. 그는 이것을 개선하고 서서히 변화할 수 있는 (slowly-evolving) 추세를 반영하는 모형으로서 비관측요인모형 (unobserved components model, UC모형; structural time series model, 구조적 시계열 모형)을 제안하였다. 먼저, UC모형은 시계열 분해 (decomposition) 기법에 그 토대를 두고 있다. 시계열이 가지고 있는 다양한 특징을 비관측적인 여러 요인들로 분해한다. 두번째 특징은, 서로 가까운 시계열들의 상관성을 잘 반영하기 위하여 국소 모형 (local model)의 입장을 취한다는 것이다. 즉, 모든 시계열 관측치에 동일한 가중치를 주는 결정적 추세는 바람직하지 못하다는 입장이다. UC모형에 관한 전반적인 소개는 Harvey (2004)를 참고하기 바란다. 본 논문은 모형의 우수성에 대해서 비교적 덜 알려져 있는 UC모형을 소개하고, 한국의 국내총생산 (gross domestic product) 시계열 자료분석에 적용하여 본다. UC모형의 유연성 (flexibility)을 이용하여 다양한 형태로 국내총생산 자료의 모형화를 시도하고, 지수평활법 및 박스-젠킨스의 ARIMA모형과 예측력을 비교한다.

<sup>†</sup> 이 논문은 2009년도 정부 (교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임 (과제번호: 2009-0071493).

<sup>1</sup> 교신저자: (156-756) 서울시 동작구 흑석동 221번지, 중앙대학교 응용통계학과, 부교수.

E-mail: bcseong@cau.ac.kr

<sup>2</sup> (156-756) 서울시 동작구 흑석동 221번지, 중앙대학교 통계학과, 석사과정.

## 2. UC모형

UC모형의 일반화된 형태는 단변량 시계열  $y_t$ 에 대하여 다음과 같이 표현된다;

$$y_t = \mu_t + \gamma_t + \psi_t + r_t + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^m \delta_j x_{jt} + \varepsilon_t, \varepsilon_t \sim iid N(0, \sigma_\varepsilon^2). \quad (2.1)$$

여기서,  $\mu_t$ 는 추세요인 (trend component),  $\gamma_t$ 는 계절요인 (seasonal component),  $\psi_t$ 는 순환요인 (cyclical component),  $r_t$ 는 자기회귀요인 (autoregressive component),  $\varepsilon_t$ 는 불규칙요인 (irregular component)을 나타낸다. 이 5가지 요인들은 모두 비관측적으로 가정되며,  $y_t$  및  $x_{jt}$ 를 통하여 추정해야 한다. 또한,  $y_t$ 의 과거값들의 영향을 나타내는  $\sum_{i=1}^p \phi_i y_{t-i}$  항 및 설명변수  $x_{jt}$  ( $j = 1, \dots, m$ )의 영향을 나타내는  $\sum_{j=1}^m \delta_j x_{jt}$  항은 전이함수모형 (transfer function model)의 형태와 유사하다고 볼 수 있다. 이와 같은 유연한 UC모형의 형태 (특히, 자기회귀요인  $r_t$ 와  $\sum_{j=1}^m \delta_j x_{jt}$  항의 포함)는, 전통적인 가법적 (additive) 시계열 분해법과는 대조적이다. 참고로, 전통적 시계열 분해법의 예로서, X12-ARIMA 및 요소분해 모형을 들 수 있으며, 이에 대한 설명은 Hwang와 Yang (2006) 및 박인찬 등 (2009)를 참고하여야 한다. UC모형의 가장 큰 특징으로 볼 수 있는 것은 추세, 계절, 순환, 자기회귀와 같은 비관측 요인들이 확률적 및 결정적 형태를 모두 포함할 수 있다는 점이며, 이것은 각 요인들과 관련된 오차 분산의 존재 유무에 따라 구분되어진다.

### 2.1. 추세요인의 모형화

추세요인을 모형화하는 방법은 크게 두가지 형태로 나누어진다; (i) 랜덤워크 모형 (random walk model)과 (ii) 국소선형 추세모형 (locally linear trend model, LLT모형). 먼저, 절편없는 (without drift) 랜덤워크 모형은 추세요인을 다음과 같이 표현한다;

$$\mu_t = \mu_{t-1} + \eta_t, \eta_t \sim iid N(0, \sigma_\eta^2). \quad (2.2)$$

이 모형은 시계열의 움직임이 장기간에 걸쳐 상승 또는 반복하는 (flat and slow-turning) 경우에 적합하다. 특히,  $\sigma_\eta^2 = 0$ 인 경우  $\mu_t = \mu_0$  (상수)이므로, 정상적 (stationary) 추세를 나타낸다고 할 수 있다. 참고로, 시계열  $y_t$ 가 식 (2.2)과 같은 추세를 가지고 있는 경우, 박스-젠킨스 모형인 ARIMA (0,1,1)과 유사하다고 볼 수 있다. 한편, LLT모형은 다음과 같은 2개의 식으로 표현된다;

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \eta_t \sim iid N(0, \sigma_\eta^2), \quad (2.3)$$

$$\beta_t = \beta_{t-1} + \xi_t, \xi_t \sim iid N(0, \sigma_\xi^2). \quad (2.4)$$

여기서,  $\eta_t$  및  $\xi_t$ 는 서로 독립이며,  $\mu_t$ 는 확률적 수준 (stochastic level),  $\beta_t$ 는 확률적 기울기 (stochastic slope)를 나타낸다. 이와 같은 확률적 요인을 사용하면 시간에 따라서 서서히 변하는 기울기 (slowly-evolving slope)를 가지는 추세를 표현할 수 있다. LLT모형은 ARIMA (0,2,2)와 유사하다고 볼 수 있다.  $\sigma_\xi^2 = 0$ 인 경우  $\beta_t = \beta_0$ 의 절편을 가지는 랜덤워크 모형으로 볼 수 있으며,  $\sigma_\xi^2 = \sigma_\eta^2 = 0$ 인 경우 결정적 선형 추세 (deterministic linear trend) 모형인  $y_t = \mu_0 + \beta_0 t + \varepsilon_t$ 의 형태를 따른다고 볼 수 있다. 즉, 확률적 수준 및 기울기의 분산에 의하여 다양한 시계열의 움직임을 모형화할 수 있다.

### 2.2. 순환요인의 모형화

순환요인의 모형화는 일반적으로 삼각함수를 이용한 방법이 널리 사용된다. 요인의 형태가 결정적이나 확률적이냐에 따라 다음의 2가지 형태로 모형화가 이루어진다. 결정적 순환모형 (deterministic

cyclical model)은 진동수  $\lambda$  ( $0 < \lambda < \pi$ )를 이용하여 다음과 같이 표현된다;

$$\psi_t = \alpha \cos(\lambda t) + \beta \sin(\lambda t). \quad (2.5)$$

따라서,  $t$ 가 연속적인 경우  $\psi_t$ 는  $2\pi/\lambda$ 의 주기를 가지는 주기함수로서 진폭(amplitude)은  $(\alpha^2 + \beta^2)^{1/2}$ , 위상각(phase angle)은  $\tan^{-1}(\beta/\alpha)$ 이다. 실제 시계열 분석에서는, 시계열의 복잡한 움직임을 모형화하기 위하여 한 개 이상의  $\psi_t$ 들의 합으로서 순환요인을 모형화하기도 한다. 확률적 순환모형(stochastic cyclical model)은 확률적 움직임을 나타내기 위하여 고정된 모수  $\alpha$ 와  $\beta$  대신에 추가적 요인(extra component)  $\psi_t^*$  및 감쇠치(damping factor)  $\rho$  ( $0 \leq \rho \leq 1$ )를 사용하여 다음과 같이 표현된다;

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda & \sin \lambda \\ -\sin \lambda & \cos \lambda \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \nu_t \\ \nu_t^* \end{bmatrix}. \quad (2.6)$$

여기서,  $\nu_t \sim iid N(0, \sigma_\nu^2)$ ,  $\nu_t^* \sim iid N(0, \sigma_{\nu^*}^2)$ 이며 서로 독립이다. 이 모형은 추가적인 다른 모수들을 사용하지 않고서도 복잡한 시계열의 움직임을 잘 모형화할 수 있음이 알려져 있다.  $\rho = 1$ 인 경우는 비정상적(nonstationary) 순환을 나타내며,  $\sigma_\nu^2 = 0$ 인 경우 식 (2.6)은 (2.5)의 특별한 형태임을 보일 수 있다.

### 2.3. 계절요인의 모형화

계절요인도 순환요인과 유사하게 삼각함수를 이용하여 모형화할 수도 있으나, 본 논문에서는 이해의 편의를 위하여 가변수(dummy variable)를 사용한 모형화를 설명하기로 한다. 전자의 방법을 위해서는 Harvey (1989)를 참고하여야. 결정적 가변수를 이용한 계절요인은, 계절주기(seasonal period)  $s$ 에 대하여  $\gamma_j$  ( $j = 1, \dots, s$ )이 계절  $j$ 에서의 고정된(fixed) 계절효과를 나타낸다고 할 때 다음의 제약식으로 표현된다;

$$\sum_{j=1}^s \gamma_j = 0. \quad (2.7)$$

따라서,  $t$ -시점에서의 계절효과  $\gamma_t$ 는 다음과 같이 정의될 수 있다;

$$\gamma_t = \begin{cases} \gamma_i & \text{만약 } t = i, i + s, i + 2s, \dots \text{ 이고 } i = 1, \dots, s-1 \\ -\sum_{j=1}^{s-1} \gamma_j & \text{만약 } t = s, 2s, 3s, \dots \end{cases} \quad (2.8)$$

또한, 가설  $H_0 : \gamma_1 = \gamma_2 = \dots = \gamma_{s-1} = 0$ 의 검정을 통하여 시계열  $y_t$ 에 존재하는 계절성의 유무를 검정할 수 있을 것이다. 확률적 가변수를 이용한 계절요인은 식 (2.7)의 우변에 오차항을 도입하여 다음과 같이 표현될 수 있다;

$$\sum_{i=0}^{s-1} \gamma_{t-i} = \omega_t, \omega_t \sim iid N(0, \sigma_\omega^2). \quad (2.9)$$

즉, 확률적 계절요인들의 합은 평균이 0이며  $\sigma_\omega^2$ 의 크기에 따라서 계절요인의 변화 움직임을 모형화하는 것이다. 결정적 가변수를 이용한 계절요인은 식 (2.9)에서  $\sigma_\omega^2 = 0$ 인 경우로 간주할 수 있다.

### 2.4. 자기회귀요인의 모형화

자기회귀요인은 순환요인의 대안으로서 흔히 사용되며 다음과 같이 AR (1)의 형태가 주로 사용된다;

$$r_t = \rho r_{t-1} + v_t, v_t \sim iid N(0, \sigma_v^2). \quad (2.10)$$

단,  $0 < \rho < 1$ 이다. 식 (2.10)의 간단한 형태에도 불구하고 많은 경영경제 시계열에 존재하는 순환을 잘 설명한다고 알려져 있다.

## 2.5. UC모형의 추정방법

UC모형은 가우시안 상태공간모형 (Gaussian state space model)의 특별한 경우로 알려져 있다. 따라서, UC모형은 상태공간모형으로 표현되고 칼만필터 (Kalman filter)를 통하여 추정된다. 이를 위하여, Harvey (1989)와 Durbin과 Koopman (2001)를 참고하여야. SAS에서는 'Proc UCM' 프로시저에서 칼만필터를 사용하여 UC모형의 다양한 추정결과를 제공한다. 특히, 각 요인들과 관련된 오차의 분산의 추정치와 t-검정, 각 요인들의 유의성에 대한 카이제곱검정 등은 최적의 모형을 찾는 데 필수적이라고 하겠다. 추정된 모형은 일반적인 시계열분석 절차에서처럼 백색잡음 (white noise) 검정 및 정규성 검정을 통하여 잔차분석이 이루어진다.

## 3. 한국 국내총생산의 UC 모형화

본 연구에 사용된 자료는 1970년 1/4분기부터 2010년 4/4분기까지의 분기별 한국의 실질 (real) 국내총생산 시계열 자료로서 계절조정 (seasonal adjustment)이 되지 않은 원자료이며, 한국은행 경제통계시스템 (<http://ecos.bok.or.kr>)에서 이용가능하다. 1970년 1/4분기부터 2008년 4/4분기까지 총 156개의 자료는 시계열모형의 모형화 (modeling)를 위하여 사용하였으며, 이후의 자료는 모형들의 예측력 비교를 위하여 사용하였다. 실질 국내총생산은 다른 대부분의 경제경영시계열 자료에서 나타나는 (점차로 증가하면서 그 변동폭이 점차로 커지는) 형태를 가지므로 분산안정화 변환 (variance stabilization)를 위하여 로그변환하였다. 그림 3.1은 2008년 4/4분기까지의 로그변환된 실질 국내총생산 변수  $\log(\text{real GDP})$ 의 시계열도표 (time series plot)이며, 전반적으로 증가하는 추세와 계절성을 뚜렷하게 가지고 있음을 알 수 있다. 또한, 2번의 뚜렷한 변화점을 관찰할 수 있으며, 이것은 1979년 1/4분기의 제2차 석유파동과 1997년 12월 IMF사태를 그 원인으로 간주하였다. 이 2가지 커다란 사회적 변화가 실제로 국내총생산 시계열에 미친 영향은 각각 1980년 1/4분기 ( $T_1$ ) 및 1998년 1/4분기 ( $T_2$ )로 확인할 수 있다. 이러한 변화점들의 영향을 정확하게 반영하기 위하여 실질 국내총생산의 모형화에서 두 시점  $T_1$ 과  $T_2$ 에서 계단함수 (step function) 형태의 개입변수 (intervention variables)  $x_{1t} = I(t \geq 1980Q1)$ 과  $x_{2t} = I(t \geq 1998Q1)$ 를 고려하였다. 개입분석 (intervention analysis)를 이용한 시계열 자료분석에 관하여는 최성관 (2000)을 참고하여야. 이후 모든 분석에서는 SAS/ETS 9.1을 이용하였다.

### 3.1. UC모형의 적합

일반적으로 UC모형의 적합은, 식 (2.1)에서 추세요인 ( $\mu_t$ ), 계절요인 ( $\gamma_t$ ) 및 불규칙 요인 ( $\varepsilon_t$ )만을 가지는 다음의 기본구조 모형 (basic structural model)에서부터 시작한다. 따라서, 개입변수  $x_{1t}$ 과  $x_{2t}$ 를 고려한 다음의 모형을 고려할 수 있다;

$$BSM1: y_t = \mu_t + \gamma_t + \delta_1 x_{1t} + \delta_2 x_{2t} + \varepsilon_t. \quad (3.1)$$

여기서,  $\mu_t$ 는 LLT모형인 식 (2.3)과 (2.4)를 따른다. 이 모형에 대한 SAS 적합 결과는 표 3.1과 같다. 따라서, 식 (3.1)에서 불규칙 요인 및 기울기 요인에 대한 오차분산의 추정치는 0으로 간주하는 것이 타당하므로, 다음의 모형을 고려할 수 있다;

$$BSM2: y_t = \mu_t + \gamma_t + \delta_1 x_{1t} + \delta_2 x_{2t}, \mu_t = \mu_{t-1} + \beta_0 + \eta_t. \quad (3.2)$$

표 3.2는 식 (3.2)의 추정 결과이다. 또한, 각 요인들의 유의성에 대한 카이제곱검정 결과 모두 유의하였고, 잔차의 자기상관 (ACF) 도표, 히스토그램 및 정규확률 (Q-Q) 그림을 검토한 결과 오차의 백색잡음 및 정규성 가정을 모두 만족하였다. 참고로, 이를 위한 SAS 프로그램은 표 3.3과 같다.

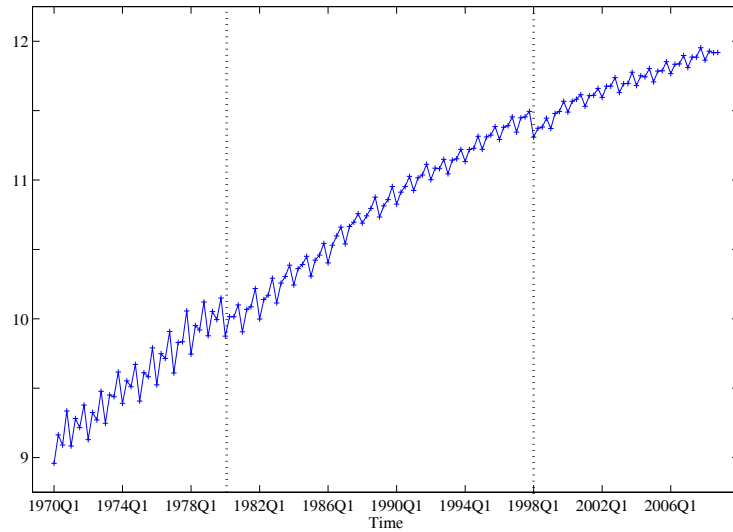


그림 3.1 1970년 1/4분기부터 2008년 4/4분기까지 log (real GDP)의 시계열도표

표 3.1 모형 (3.1)에 대한 추정결과

요인	모수	추정치	표준오차	t-값	p-값
불규칙 $\varepsilon_t$		$1.8664 \times 10^{-10}$	$1.1886 \times 10^{-7}$	0.00	0.9987
수준 $\mu_t$	오차분산	$2.3733 \times 10^{-4}$	$4.0620 \times 10^{-5}$	5.84	< .0001
기울기 $\beta_t$		$6.8461 \times 10^{-7}$	$6.4408 \times 10^{-7}$	1.06	0.2878
계절 $\gamma_t$		$3.4790 \times 10^{-5}$	$9.3868 \times 10^{-6}$	3.71	0.0002
$x_1$	계수	-0.0512	0.0192	-2.67	0.0077
$x_2$		-0.1024	0.0192	-5.34	< .0001

표 3.2 모형 (3.2)에 대한 추정결과

요인	모수	추정치	표준오차	t-값	p-값
수준 $\mu_t$	오차분산	$2.8209 \times 10^{-4}$	$4.3730 \times 10^{-5}$	6.45	<.0001
계절 $\gamma_t$		$3.2340 \times 10^{-5}$	$8.7054 \times 10^{-6}$	3.72	0.0002
$x_1$	계수	-0.0462	0.0202	-2.29	0.0218
$x_2$		-0.1049	0.0202	-5.20	<.0001
$\beta_0$	상수	0.0190	0.0014	13.57	<.0001

모형 (3.2)에 순환요인 ( $\psi_t$ )을 추가한 다음의 모형도 고려할 수 있다;

$$BSM3 : y_t = \mu_t + \gamma_t + \psi_t + \delta_1 x_{1t} + \delta_2 x_{2t}, \mu_t = \mu_{t-1} + \beta_0 + \eta_t. \tag{3.3}$$

그러나, 모형을 추정된 결과 순환요인의 오차분산 및 주기가 유의하지 않았기 때문에 요인을 포함시키지 않는 것이 바람직하다. 자기회귀요인 및  $y_t$ 의 과거값들을 독립변수로 추가한 모형도 고려할 수 있으나, 역시 유의하지 않았다. 표 3.4는 고려할 수 있는 UC모형들의 정보량 기준 (information criteria), 제곱근 MSE (root mean squared error, RMSE) 및 MAPE (mean absolute percentage error)의 값을 보여준다. BSM11은 BSM1에 결정적 기울기 ( $\sigma_\xi^2 = 0$ )만을 반영한 UC모형이며, BSM12는 BSM1에서 불규칙요인을 제거한 모형이다. BSM3은 식 (3.3)에서 확률적 순환요인을 고려한 것이고, BSM31은 결

표 3.3 BSM2의 적합을 위한 SAS 프로그램

```

ods html;
ods graphics on;
Proc UCM data=gdp;
id qrt interval=quarter;
model lngdp=x1 x2;
irregular var=0 noest print=smooth;
level print=smooth;
slope var=0 noest print=smooth;
season length=4 type=dummy print=smooth;
estimate plot= (residual normal acf);
forecast lead=8 plot= (forecasts decomp) outfor=result1;
run;
ods graphics off;
ods html close;

```

정적 순환요인을 고려한 것이다. RMSE 및 MAPE 측면에서는 BSM3가, AIC (Akaike, 1973) 및 BIC (Schwarz, 1978) 값으로 판단할 때 BSM12가 최적의 모형이라고 보여지나, 각 요인의 오차분산의 유의성 검토 결과에 의하여, 최적의 모형으로 보이는 BSM2를 사용하여 다음 절에서 그 예측력을 비교한다.

표 3.4 고려할 수 있는 UC모형들

UC모형의 종류	AIC	BIC	RMSE	MAPE
BSM1	-661.82	-628.28	0.0176	0.1014
BSM11 ( $\sigma_{\xi}^2 = 0$ )	-654.89	-624.39	0.0205	0.1281
BSM12 ( $\sigma_{\varepsilon}^2 = 0$ )	-663.82	-633.33	0.0176	0.1014
BSM2 ( $\sigma_{\xi}^2 = \sigma_{\varepsilon}^2 = 0$ )	-656.89	-629.44	0.0205	0.1281
BSM3	-661.37	-624.77	0.0172	0.0993
BSM31 ( $\sigma_{\nu}^2 = 0$ )	-652.89	-619.34	0.0205	0.1281

### 3.2. 예측력 비교

본 논문에서는 UC모형의 예측력의 우수성을 보이기 위하여, 일반적으로 널리 사용되는 단변량 시계열 분석을 위한 모형인 지수평활법과 ARIMA모형과 비교한다. 예측력을 비교하기 위해서는 다음과 같은 방법을 사용하였다. 먼저, 표본기간을 2008년 4/4분기까지로만 잡고 각 시계열모형을 모형화(modeling)하고 이후의 관측치 8개(2009년 1/4분기부터 2010년 4/4분기)를 예측하였다. 이를  $n = 1, \dots, 8$ 에 대하여  $n$ -시차 후의 예측( $n$ -step ahead forecast)이라고 표현하자. 그리고, 8개의 참값과 예측치에 의하여 예측오차(forecast error)를 계산하였다. 다음 단계로, 표본기간을 1사분기 늘려서 2009년 1/4분기까지로 잡고 다시 각 시계열모형을 모형화하고 이후의 관측치 7개(2009년 2/4분기부터 2010년 4/4분기)에 대한  $n$ -시차 후의 예측오차( $n = 1, \dots, 7$ )를 계산하였다. 이와 같은 방법을 계속 사용하여, 표본기간으로 2010년 3/4분기까지의 자료를 이용해서 이후의 관측치 1개에 대한 예측오차까지 계산할 수 있다.

UC모형의 예측력을 평가하기 위하여, 우선 3.1절에서의 BSM2를 사용하였고 비교를 위하여 Winters의 가법계절지수평활법과 박스-젠킨스의 ARIMA모형을 고려하였다. Winters의 가법계절지수평활법에 관한 구체적 적용 방법은 Song 등(2008)을 참고하여야. 모형의 특성상, 실질 국내총생산에 나타난 2번의 변화점은 UC모형 및 ARIMA모형에만 반영되었다. 지수평활법에서 사용되어지는 평활상수들은 잔차제곱합을 최소화 하는 값들을 사용하였으며, ARIMA모형은 계절성 반영 및 잔차의 포맷투

(Portmanteau) 검정까지 고려하여  $ARIMA(1, 1, 1) \times (0, 1, 1)_{S=4}$ 을 사용하였다. 참고로, 그림 3.2는 UC모형의 적합결과를 살펴보기 위하여, 2002년부터 2008년 이후 8-시차까지 예측치 및 신뢰구간을 표시한 것이다.

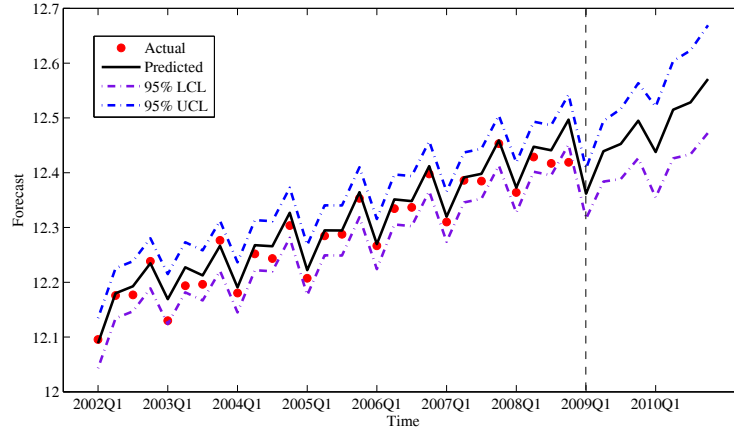


그림 3.2 BSM2에 의한 2002년 1/4분기부터 2010년 4/4분기까지의 예측치 및 신뢰구간

표 3.5 3가지 모형 (UC모형, ARIMA, 지수평활법)의 예측 시차별 예측력 (RMSE) 계산

$n-lags$	1	2	3	4	5	6	7	8	total
UCM	0.0181	0.0205	0.0220	0.0210	0.0255	0.0287	0.0443	0.0454	0.0246
ARIMA	0.0173	0.0246	0.0354	0.0437	0.0545	0.0644	0.0724	0.0695	0.0419
Winters	0.0176	0.0214	0.0246	0.0259	0.0287	0.0347	0.0367	0.0213	0.0251

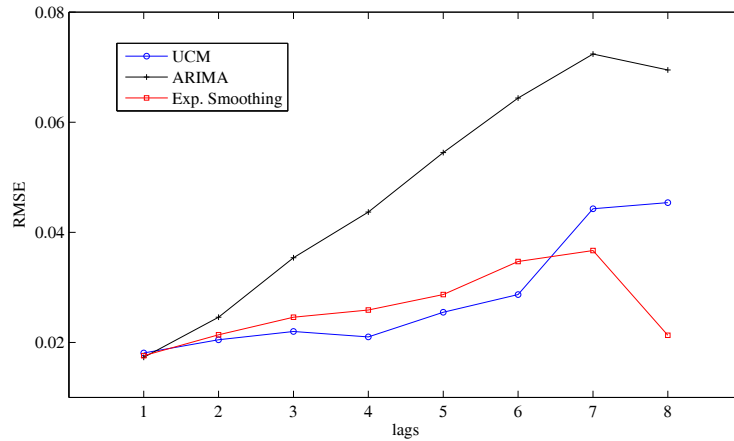


그림 3.3 3가지 모형 (UC모형, ARIMA, 지수평활법)의 예측력 비교

표 3.5는 3가지 모형 (UC모형, 지수평활법, ARIMA모형)에 대하여, 동일한 예측 시차별로 예측오차

의 RMSE를 계산한 것이다. 우선, 모든 시차에 대한 RMSE를 모두 더한 값으로 평가할 경우, ‘UC모형 > 지수평활법 > ARIMA모형’ 순서로 예측력이 우수하였다. 예측 시차별로 살펴보면, 2-시차부터 6-시차까지의 모든 시차에서 UC모형의 예측이 우수하였으며, ARIMA모형이 1-시차에만 가장 예측력이 좋다는 것은 흥미로운 결과이다. 7-시차 및 8-시차에서는 Winters의 가법계절지수평활법이 UC모형보다 예측력이 우수하다는 점을 보이고 있으나, 사용된 관측치가 각각 2개와 1개이므로 예측력을 비교하기에는 무리가 있다고 보여진다. 다만, 표 3.5 및 그림 3.3를 통하여 볼 때, ARIMA모형의 예측력은 장기적으로 갈수록 그 정확성이 다른 모형에 비하여 빠르게 떨어지는 점을 관찰할 수 있다.

#### 4. 결론

Harvey (1989) 이후, UC모형은 단변량 및 다변량 시계열 분석을 위한 많은 응용 분석법들로 발전되어져 왔다. 그러나, 실무 분석가들은 시계열 자료를 모형화하고 예측함에 있어 여전히 지수평활법 및 박스-젠킨스의 ARIMA 모형과 같은 전통적 분석법에 의존하고 있다. 추세, 순환, 계절 및 불규칙 요인과 같은 비관측 요인을 이용함으로써 UC모형은 ARIMA 모형의 유연성 및 지수평활법의 해석의 편리함을 모두 갖추고 있다고 알려져 있다. 본 논문에서는 그 우수성에 비하여 비교적 덜 알려져 있는 UC모형을 소개하고 한국의 국내총생산 시계열 자료에 적용한다. 다양한 형태의 모형화를 시도하며, 지수평활법 및 ARIMA모형과 예측력을 비교한다. 이를 토대로, 향후 다변량 UC모형의 적용에 관한 연구의 가능성을 보여준다.

#### 참고문헌

- 박인찬, 권오진, 김태윤 (2009). 시계열 모형을 이용한 주가지수 방향성 예측. <한국데이터정보과학회지>, **20**, 991-998.
- 최성관 (2000). 시계열모형을 이용한 선거개입의 경제적 영향분석. <한국데이터정보과학회지>, **11**, 257-267.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, 267-281, Akademiai Kiado, Budapest.
- Diebold, F. X. and Senhadji, S. A. (1996). *Deterministic vs stochastic trends in U.S. GNP, yet again*, Working Paper No. 5481, National Bureau of Economic Research.
- Durbin, J. and Koopman, S. J. (2001). *Time series analysis by state space methods*, Oxford University Press, Oxford.
- Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Cambridge, UK.
- Harvey, A. C. (1997). Trends, cycles and autoregression. *Econometric Journal*, **107**, 192-201.
- Harvey, A. C. (2004). Forecasting with unobserved components time series models. In *Handbook of Economic Forecasting*, edited by G. Elliott, C. W. J. Granger and A. Timmermann, 327-412, Elsevier, Amsterdam.
- Hwang, S. Y. and Yang, S. K. (2006). A refinement of point forecast using dependency structure in irregular component of BOK-X12-ARIMA. *Journal of the Korean Data & Information Science Society*, **17**, 141-147.
- Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica*, **57**, 1361-1401.
- Perron, P. and Wada, T. (2005). *Trends and cycles: A new approach and explanations of some old puzzles*, Manuscript No. 252, Department of Economics, Boston University.
- Rao, B. B. (2010). Deterministic and stochastic trends in the time series models: A guide for the applied economist. *Applied Economics*, **42**, 2193-2202.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.
- Song, P. J., Um, H. J. and Kim, J. T. (2008). Forecasting of stream qualities at Gumi industrial complex by Winters' exponential smoothing. *Journal of the Korean Data & Information Science Society*, **19**, 1133-1140.



# Analysis of Korean GDP by unobserved components model<sup>†</sup>

Byeongchan Seong<sup>1</sup> · Seung-Kyung Lee<sup>2</sup>

<sup>1</sup>Department of Applied Statistics, Chung-Ang University

<sup>2</sup>Department of Statistics, Chung-Ang University

Received 13 July 2011, revised 7 August 2011, accepted 15 August 2011

## Abstract

Since Harvey (1989), many approaches for applying unobserved components (UC) models to both univariate and multivariate time series analysis have been developed. However, practitioners still tend to use traditional methods such as exponential smoothing or ARIMA models for modeling and predicting time series data. It is well known that the UC model combines the flexibility of ARIMA models and the easy interpretability of exponential smoothing models by using unobserved components such as trend, cycle, season, and irregular components. This study reviews the UC model and compares its relative performances with those of the other models in modeling and predicting the real gross domestic products (GDP) in Korea. We conclude that the optimal model is the UC model on basis of root mean squared error.

*Keywords:* State space model, stochastic trends, structural time series model.

---

<sup>†</sup> This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant number: 2009-0071493).

<sup>1</sup> Corresponding author: Associate professor, Department of Applied Statistics, Chung-Ang University, 221, Heukseok-dong, Dongjak-gu, Seoul 156-756, Korea. E-mail: bcseong@cau.ac.kr

<sup>2</sup> Graduate student, Department of Statistics, Chung-Ang University, 221, Heukseok-dong, Dongjak-gu, Seoul 156-756, Korea.